Ahsanullah University of Science and Technology

Department of Computer Science and Engineering

Soft Computing Lab
CSE4238

Assignment - 2

Deep Neural Network

Submitted to

Mr. Nibir Chandra Mandal
Lecturer, CSE, AUST

Sanzana Karim Lora
Lecturer, CSE, AUST

Submitted by

Tashfiq Ahmed   **17.01.04.038**

September 30, 2021

Contents

1   Dataset 1

1.1   Experiment 1

- Batch Size = 20

- Iteration = 20000

- Epoch = 21

- Input Dimension = 28*28

- Learning Rate = 0.01

- Accuracy = 9.75%

- Loss = 2.30

As the number hidden node was very less, and also the batch size was not so much means ot was not good enough in proportion to the dataset that is why the result was not so satisfactory as well. We can see that from the accuracy graph and loss graph as well.
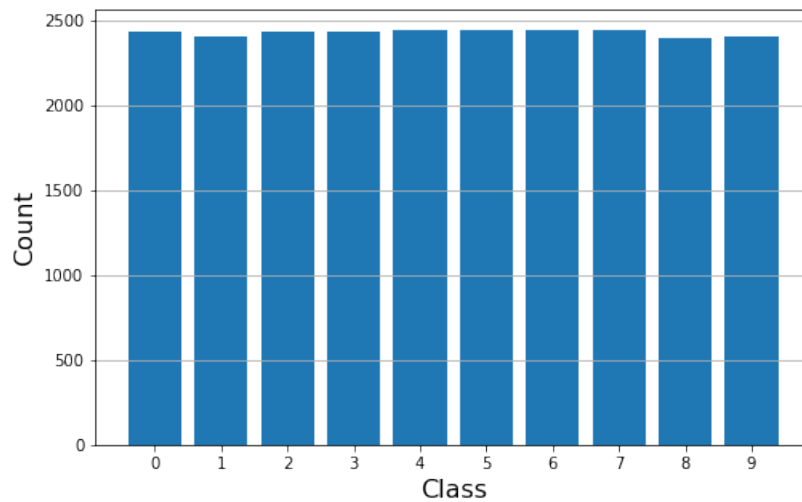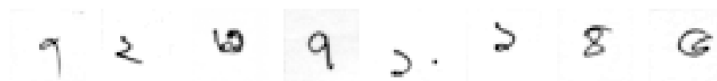

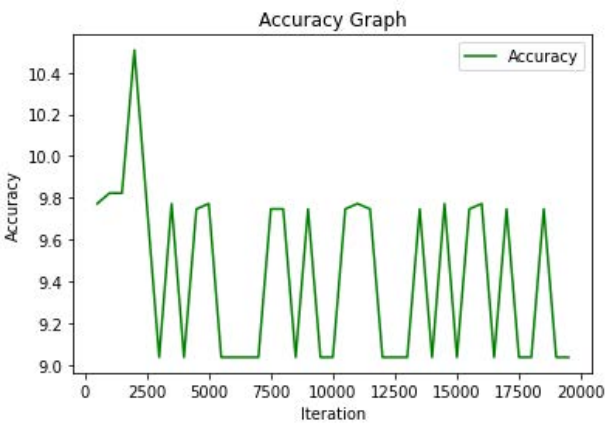Figure 1: Data Visualization

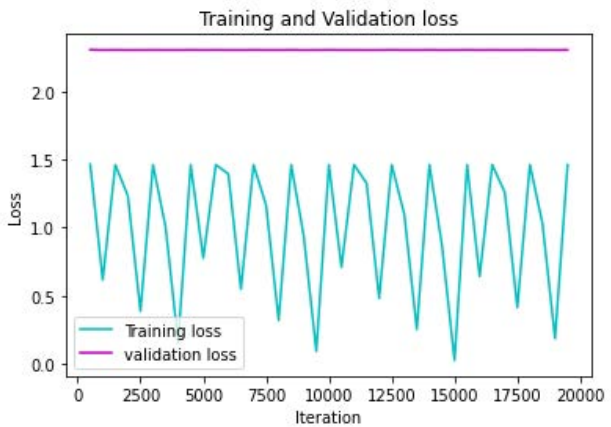Figure 2: Dataset

Figure 3: Accuracy Graph
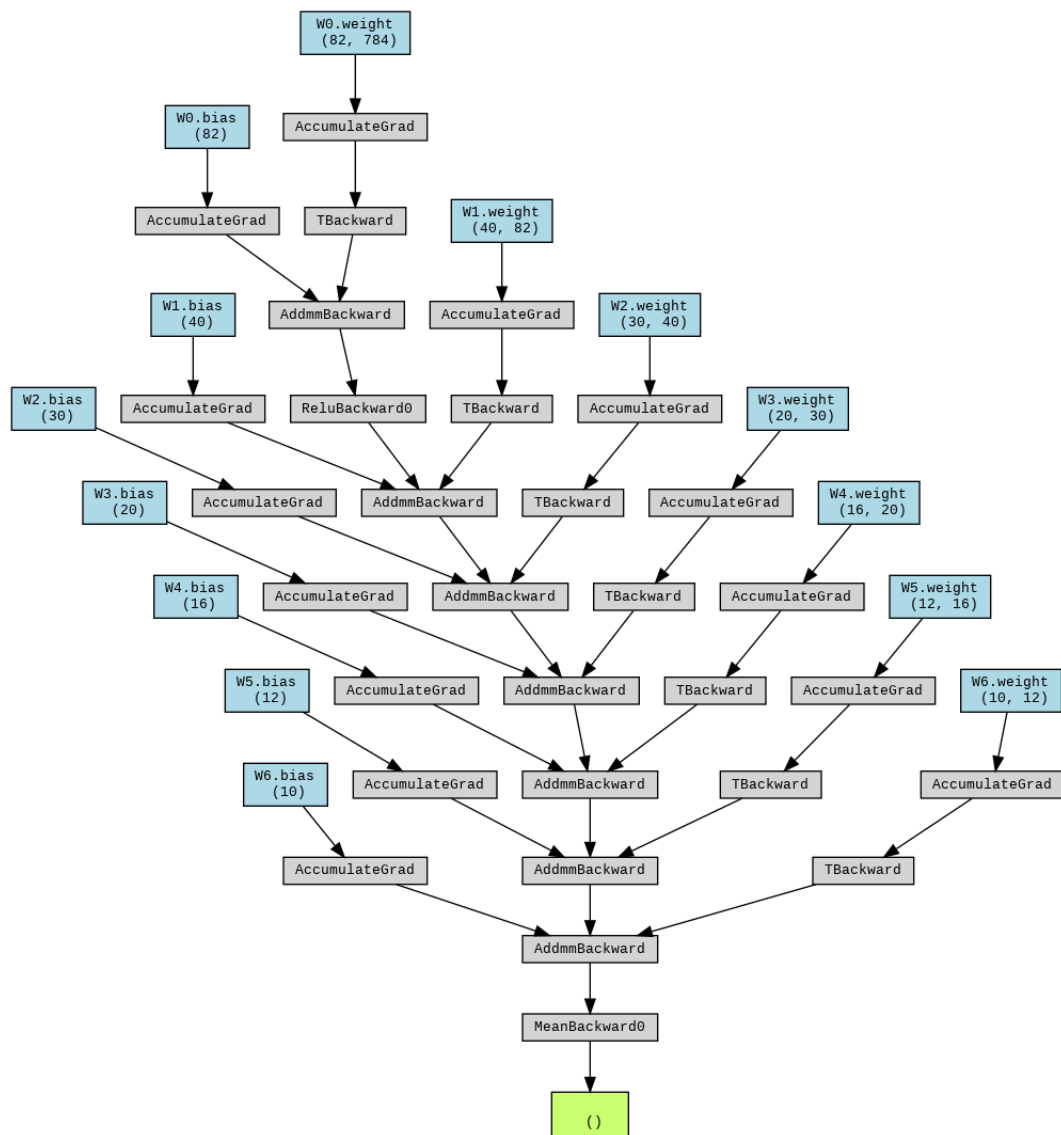

Figure 4: Loss Graph

Figure 5: Model Visualization

1.2 Experiment 2, Trial 2

- Batch Size = 32

- Iteration = 125000

- Epoch = 203

- Input Dimension = 28*28

- Learning Rate = 0.0001

- Accuracy = 91.21%

- Loss = 0.28

After increasing the number of iterations and decreasing the learning rate the result came out quite good. As the learning rate was much less compared to experiment 1, it could converge more precisely. More importantly as we increased the number of nodes in the hidden layer the change in the accuracy and loss was really impressive. The early layers detect low level features (edges and corners), and later layers successively detect higher level features(classify an object). In our problem, as we had to classify bangla digits[0-9] our hidden layers converged to 10. The activation function was ReLU in all the cases and in optimizer Adam, Adamax and SGD was used but the result was improved significantly after using SGD optimizer.
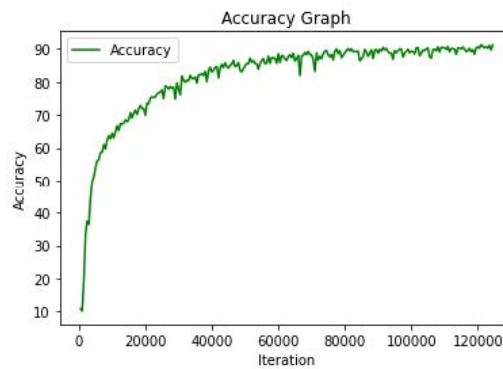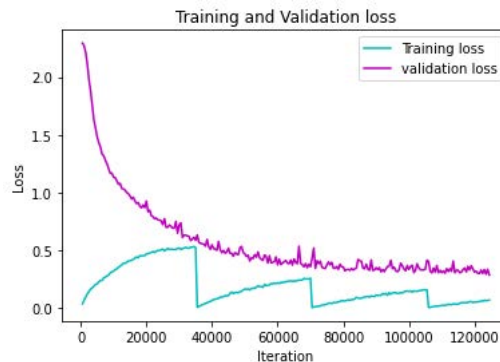

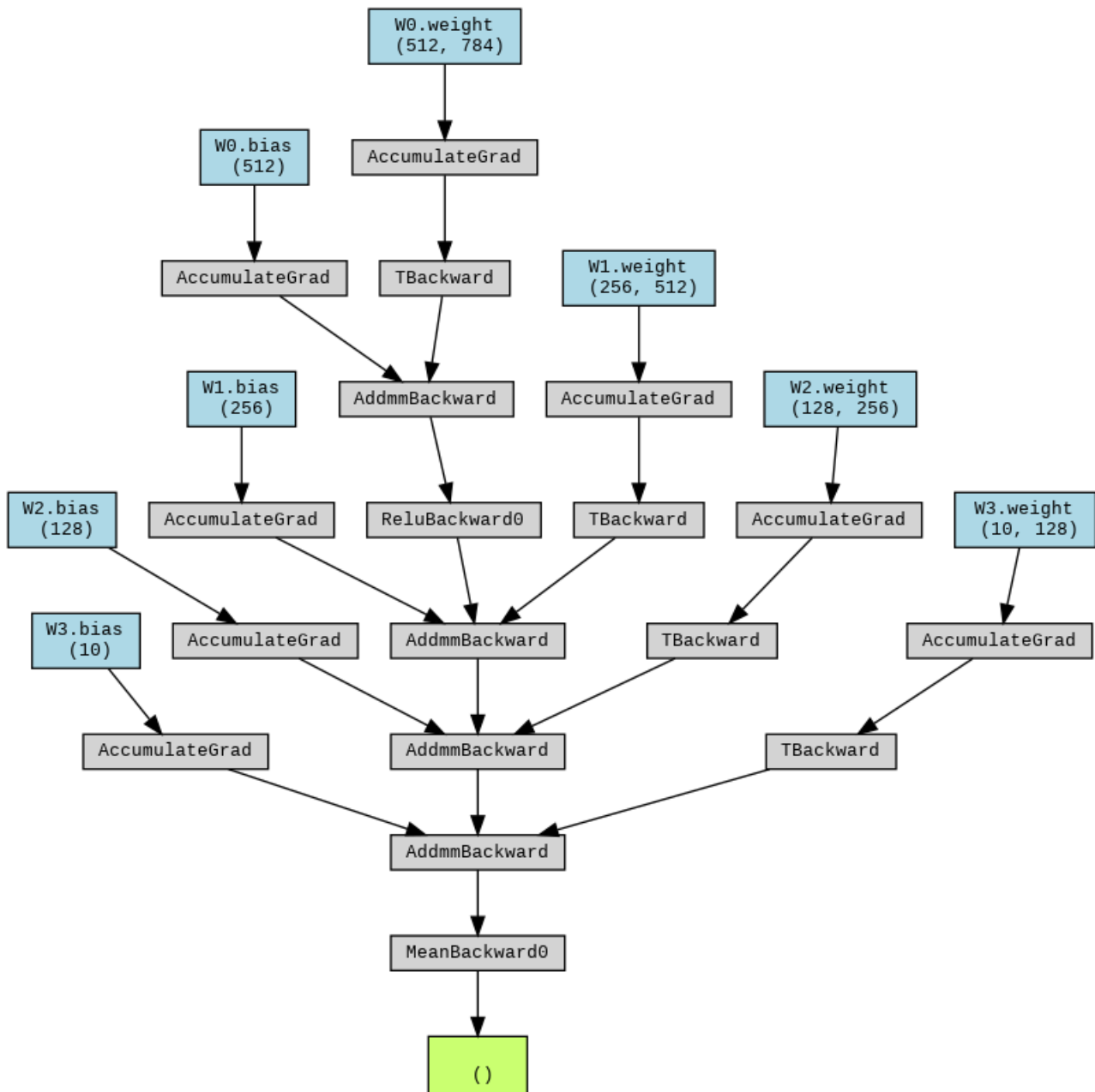Figure 6: Accuracy Graph

Figure 7: Loss Graph

Figure 8: Model Visualization

2 Comparison between experiments

In this experiment the task was to build a multiclass classifier using deep neural network that would recognize bengali digits[0-9]. For this experiment a dataset was provided containing 19702 images and training-testing ratio was 80/20. For, first experiment the hyper-parameters were pre defined in the assignment provided to us. Next our task was to tune them so that we could generate a better model in terms of accuracy and loss.

No	Accuracy	Loss	Batch Size	Number of Iterations	Epoch	Learning Rate	Optimizer	Hidden Nodes	Hidden Layers	Activation Function
1	9.75%	2.30	20	20000	21	0.01	SGD	200	6	ReLU
2	45.96%	2.30	200	20000	204	0.01	SGD	1800	3	ReLU
3	91.21%	0.28	32	125000	203	0.001	Adam	896	3	ReLU

3 Dataset 2

3.1 Experiment 3, Model 1

For dataset 2, the hyper paramters we used for experiment 1 was matched with the data precisely. The accuracy was 86.5% and the valid loss was 0.377.

- Batch Size = 20

- Iteration = 20000

- Epoch = 21

- Input Dimension = 28*28
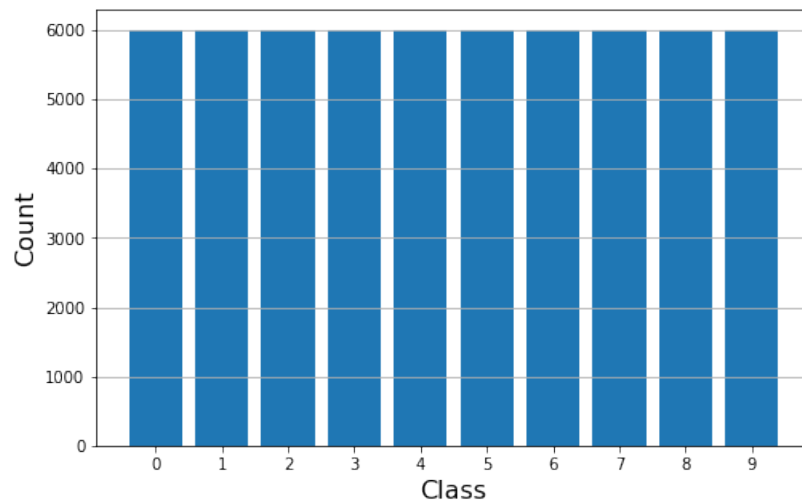
- Learning Rate = 0.01

- Accuracy = 90.23%

- Loss = 0.42
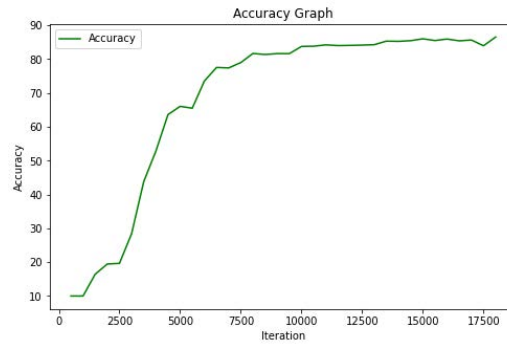

Figure 9: Data Visualization
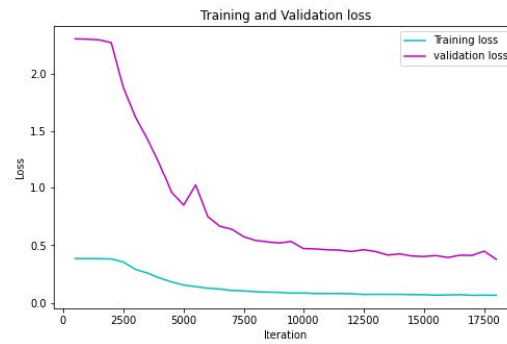

Figure 10: Dataset

Figure 11: Accuracy Graph

Figure 12: Loss Graph

Hyper-parameter: Given	Accuracy	Loss
Dataset 1	9.75%	2.30
Dataset 2	86.5%	0.37

3.2 Experiment 3, Model 2

The hyper parameters we used for experiment 2 were used here as well and the result was even better than trail 1. The accuracy was 90.23% and valid loss was 0.72.

- Batch Size = 32

- Iteration = 125000

- Epoch = 203

- Input Dimension = 28*28

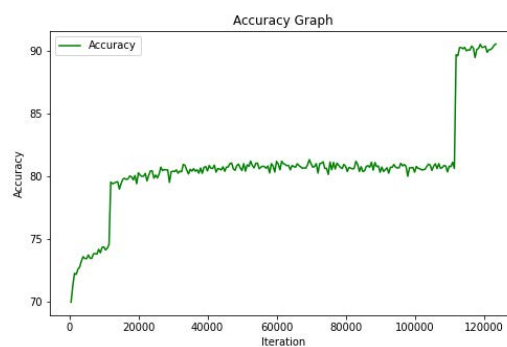- Learning Rate = 0.00001

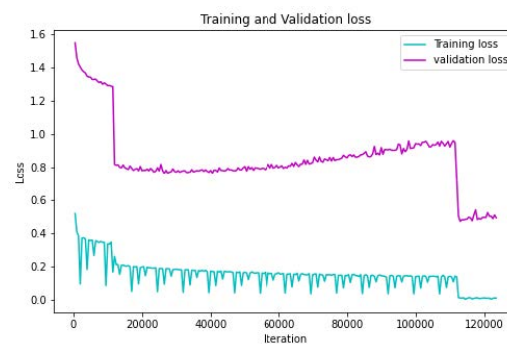- Accuracy = 90.23%

- Loss = 0.72

Figure 13: Accuracy Graph

Figure 14: Loss Graph

Hyper-parameter: Modified	Accuracy	Loss
Dataset 1	91.21%	0.28
Dataset 2	90.23%	0.72

4   Git Link

Assignment 2